

N. SALAMI

RECEIVED
FEB 9 2000
1648
USPTO MAIL ROOM

ENTERED

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/000,004

DATE: 08/02/2000
TIME: 13:46:20

Input Set : A:\Sabe-137.app
Output Set: N:\CRF3\08022000\I000004.raw

SEQUENCE LISTING

4 (1) GENERAL INFORMATION:

6 (i) APPLICANT: WEBB, Elizabeth Ann
7 MARGETTS, Mary Brigid
8 COX, John Cooper
9 FRAZER, Ian
10 MCMILLAN, Nigel Alan John
11 WILLIAMS, Mark Philip
12 MOLONEY, Margaret Bridget
13 Holland
14 EDWARDS, Stirling John
15 (ii) TITLE OF INVENTION: PAPILLOMAVIRUS POLYPROTEIN CONSTRUCTS
16 (iii) NUMBER OF SEQUENCES: 50
17 (iv) CORRESPONDENCE ADDRESS:
18 (A) ADDRESSEE: FOLEY & LARDNER
19 (B) STREET: 3000 K Street, N.W.
20 (C) CITY: Washington
21 (D) STATE: D.C.
22 (E) COUNTRY: U.S.A.
23 (F) ZIP: 20007-5109

24 (v) COMPUTER READABLE FORM:

25 (A) MEDIUM TYPE: Floppy disk
26 (B) COMPUTER: IBM PC compatible
27 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
28 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

29 (vi) CURRENT APPLICATION DATA:
30 (A) APPLICATION NUMBER: US/09/000,004
31 (B) FILING DATE: 21-Apr-1998
32 (C) CLASSIFICATION:

33 (vii) PRIOR APPLICATION DATA:
34 (A) APPLICATION NUMBER: WO PCT/AU96/00473
35 (B) FILING DATE: 26-JUL-1996

36 (A) APPLICATION NUMBER: AU PN 4439/95
37 (B) FILING DATE: 27-JUL-1995

38 (viii) ATTORNEY/AGENT INFORMATION:

39 (A) NAME: BENT, Stephen A.
40 (B) REGISTRATION NUMBER: 29,768
41 (C) REFERENCE/DOCKET NUMBER: 017227/0137

42 (ix) TELECOMMUNICATION INFORMATION:

43 (A) TELEPHONE: (202) 672-5300
44 (B) TELEFAX: (202) 672-5399

45 (2) INFORMATION FOR SEQ ID NO: 1:

46 (i) SEQUENCE CHARACTERISTICS:

47 (A) LENGTH: 30 base pairs
48 (B) TYPE: nucleic acid
49 (C) STRANDEDNESS: single
50 (D) TOPOLOGY: linear

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/000,004

DATE: 08/02/2000
TIME: 13:46:20

Input Set : A:\Sabe-137.app
Output Set: N:\CRF3\08022000\1000004.raw

65 (ii) MOLECULE TYPE: DNA (genomic)		
70 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:		30
72 GCGCCCCGGG ATGGAAAGTG CAAATGCC		
74 (2) INFORMATION FOR SEQ ID NO: 2:		
76 (i) SEQUENCE CHARACTERISTICS:		
77 (A) LENGTH: 42 base pairs		
78 (B) TYPE: nucleic acid		
79 (C) STRANDEDNESS: single		
80 (D) TOPOLOGY: linear		
82 (ii) MOLECULE TYPE: DNA (genomic)		
87 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:		42
89 GCGCTCTAGA CCATGGAAGC TTGGGTAACA TGTCTCCAT GC		
91 (2) INFORMATION FOR SEQ ID NO: 3:		
93 (i) SEQUENCE CHARACTERISTICS:		
94 (A) LENGTH: 49 base pairs		
95 (B) TYPE: nucleic acid		
96 (C) STRANDEDNESS: single		
97 (D) TOPOLOGY: linear		
99 (ii) MOLECULE TYPE: DNA (genomic)		
104 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:		49
106 GCGCTCTAGA GAGCTCGGTA CCACTAGTGG AGCACCAAAC ATTGGGAAG		
108 (2) INFORMATION FOR SEQ ID NO: 4:		
110 (i) SEQUENCE CHARACTERISTICS:		
111 (A) LENGTH: 31 base pairs		
112 (B) TYPE: nucleic acid		
113 (C) STRANDEDNESS: single		
114 (D) TOPOLOGY: linear		
116 (ii) MOLECULE TYPE: DNA (genomic)		
121 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:		31
123 GCGCAGATCT TAGGCGTAGC TGAACGTGTTA C		
125 (2) INFORMATION FOR SEQ ID NO: 5:		
127 (i) SEQUENCE CHARACTERISTICS:		
128 (A) LENGTH: 33 base pairs		
129 (B) TYPE: nucleic acid		
130 (C) STRANDEDNESS: single		
131 (D) TOPOLOGY: linear		
133 (ii) MOLECULE TYPE: DNA (genomic)		
138 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:		33
140 GCGCCCATGG GAAGTGGTGC CTGTACAAAT AGC		
142 (2) INFORMATION FOR SEQ ID NO: 6:		
144 (i) SEQUENCE CHARACTERISTICS:		
145 (A) LENGTH: 32 base pairs		
146 (B) TYPE: nucleic acid		
147 (C) STRANDEDNESS: single		
148 (D) TOPOLOGY: linear		
150 (ii) MOLECULE TYPE: DNA (genomic)		
155 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:		32
157 GCGCTCTAGA TTGCTGTGTG GTAACAATAT AG		
159 (2) INFORMATION FOR SEQ ID NO: 7:		

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/000,004

DATE: 08/02/2000
TIME: 13:46:20

Input Set : A:\Sabe-137.app
Output Set: N:\CRF3\08022000\1000004.raw

```

161      (i) SEQUENCE CHARACTERISTICS:
162          (A) LENGTH: 34 base pairs
163          (B) TYPE: nucleic acid
164          (C) STRANDEDNESS: single
165          (D) TOPOLOGY: linear
167      (ii) MOLECULE TYPE: DNA (genomic)
168          (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:           34
172          (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
174 GCGCAAGCTT CATGGAAGAC ATGTTACCCCT AAAG
176 (2) INFORMATION FOR SEQ ID NO: 8:
178      (i) SEQUENCE CHARACTERISTICS:
179          (A) LENGTH: 30 base pairs
180          (B) TYPE: nucleic acid
181          (C) STRANDEDNESS: single
182          (D) TOPOLOGY: linear
184      (ii) MOLECULE TYPE: DNA (genomic)
189          (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:           30
191 GCGCCCATGG GGTCTTCGGT GCGCAGATGG
193 (2) INFORMATION FOR SEQ ID NO: 9:
195      (i) SEQUENCE CHARACTERISTICS:
196          (A) LENGTH: 35 base pairs
197          (B) TYPE: nucleic acid
198          (C) STRANDEDNESS: single
199          (D) TOPOLOGY: linear
201      (ii) MOLECULE TYPE: DNA (genomic)
206          (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:           35
208 GCGCGAGCTC GCGGACGATT CAGGTACAGA AAATG
210 (2) INFORMATION FOR SEQ ID NO: 10:
212      (i) SEQUENCE CHARACTERISTICS:
213          (A) LENGTH: 33 base pairs
214          (B) TYPE: nucleic acid
215          (C) STRANDEDNESS: single
216          (D) TOPOLOGY: linear
218      (ii) MOLECULE TYPE: DNA (genomic)
223          (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:           33
225 GCGCGGTACC TAAAGTTCTA ACAACTGTTTC CTG
227 (2) INFORMATION FOR SEQ ID NO: 11:
229      (i) SEQUENCE CHARACTERISTICS:
230          (A) LENGTH: 32 base pairs
231          (B) TYPE: nucleic acid
232          (C) STRANDEDNESS: single
233          (D) TOPOLOGY: linear
235      (ii) MOLECULE TYPE: DNA (genomic)
240          (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:           32
242 GCGCGGTACC GAAGCAATAG CCAAGCGTTT AG
244 (2) INFORMATION FOR SEQ ID NO: 12:
246      (i) SEQUENCE CHARACTERISTICS:
247          (A) LENGTH: 33 base pairs
248          (B) TYPE: nucleic acid
249          (C) STRANDEDNESS: single

```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/000,004

DATE: 08/02/2000
TIME: 13:46:20

Input Set : A:\Sabe-137.app
Output Set: N:\CRF3\08022000\1000004.raw

250	(D) TOPOLOGY: linear	
252	(ii) MOLECULE TYPE: DNA (genomic)	
257	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:	33
259	GCGCACTAGT CAATAGGTGC AGTGACATAA ATC	
261	(2) INFORMATION FOR SEQ ID NO: 13:	
263	(i) SEQUENCE CHARACTERISTICS:	
264	(A) LENGTH: 32 base pairs	
265	(B) TYPE: nucleic acid	
266	(C) STRANDEDNESS: single	
267	(D) TOPOLOGY: linear	
269	(ii) MOLECULE TYPE: DNA (genomic)	
274	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:	32
276	GCGCTCTAGA CTAACATGTC AATTAAATGA TG	
278	(2) INFORMATION FOR SEQ ID NO: 14:	
280	(i) SEQUENCE CHARACTERISTICS:	
281	(A) LENGTH: 31 base pairs	
282	(B) TYPE: nucleic acid	
283	(C) STRANDEDNESS: single	
284	(D) TOPOLOGY: linear	
286	(ii) MOLECULE TYPE: DNA (genomic)	
291	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:	31
293	GCGCGAGCTC ATTCAATAT ATATAATCAC C	
295	(2) INFORMATION FOR SEQ ID NO: 15:	
297	(i) SEQUENCE CHARACTERISTICS:	
298	(A) LENGTH: 30 base pairs	
299	(B) TYPE: nucleic acid	
300	(C) STRANDEDNESS: single	
301	(D) TOPOLOGY: linear	
303	(ii) MOLECULE TYPE: DNA (genomic)	
308	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:	30
310	GCGCCCCGGG ATGGAAGCAA TAGCCAAGCG	
312	(2) INFORMATION FOR SEQ ID NO: 16:	
314	(i) SEQUENCE CHARACTERISTICS:	
315	(A) LENGTH: 51 base pairs	
316	(B) TYPE: nucleic acid	
317	(C) STRANDEDNESS: single	
318	(D) TOPOLOGY: linear	
320	(ii) MOLECULE TYPE: DNA (genomic)	
325	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:	51
327	GCGCTCTAGA CCATGGGTA CCGAGCTCCA ATAGGTGCAG TGACATAAAT C	
329	(2) INFORMATION FOR SEQ ID NO: 17:	
331	(i) SEQUENCE CHARACTERISTICS:	
332	(A) LENGTH: 32 base pairs	
333	(B) TYPE: nucleic acid	
334	(C) STRANDEDNESS: single	
335	(D) TOPOLOGY: linear	
337	(ii) MOLECULE TYPE: DNA (genomic)	
342	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:	32
344	GCGCTCTAGA CTAACATGTC AATTAAATGA TG	

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/000,004

DATE: 08/02/2000
TIME: 13:46:20

Input Set : A:\Sabe-137.app
Output Set: N:\CRF3\08022000\1000004.raw

346 (2) INFORMATION FOR SEQ ID NO: 18:
348 (i) SEQUENCE CHARACTERISTICS:
349 (A) LENGTH: 36 base pairs
350 (B) TYPE: nucleic acid
351 (C) STRANDEDNESS: single
352 (D) TOPOLOGY: linear
354 (ii) MOLECULE TYPE: DNA (genomic)
359 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18: 36
361 GCGCAGATCT CTCGAGATT ATATATAAT AATCAC
363 (2) INFORMATION FOR SEQ ID NO: 19:
365 (i) SEQUENCE CHARACTERISTICS:
366 (A) LENGTH: 1107 base pairs
367 (B) TYPE: nucleic acid
368 (C) STRANDEDNESS: single
369 (D) TOPOLOGY: linear
371 (ii) MOLECULE TYPE: DNA (genomic)
374 (ix) FEATURE:
375 (A) NAME/KEY: CDS
376 (B) LOCATION: 1..1104
379 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19: 48
381 ATG GAA AGT GCA AAT GCC TCC ACG TCT GCA ACG ACC ATA GAC CAG TTG
382 Met Glu Ser Ala Asn Ala Ser Thr Ser Ala Thr Thr Ile Asp Gln Leu
383 1 5 10 15
385 TGC AAG ACG TTT AAT CTA TCT ATG CAT ACG TTG CAA ATT AAT TGT GTG
386 Cys Lys Thr Phe Asn Leu Ser Met His Thr Leu Gln Ile Asn Cys Val
387 20 25 30 36
389 TTT TGC AAG AAT GCA CTG ACC ACA GCA GAG ATT TAT TCA TAT GCA TAT
390 Phe Cys Lys Asn Ala Leu Thr Thr Ala Glu Ile Tyr Ser Tyr Ala Tyr
391 35 40 45 51
393 AAA CAC CTA AAG GTC CTG TTT CGA GGC GGC TAT CCA TAT GCA GCC TGC
394 Lys His Leu Lys Val Leu Phe Arg Gly Gly Tyr Pro Tyr Ala Ala Cys
395 50 55 60 65
397 GCG TGC TGC CTA GAA TTT CAT GGA AAA ATA AAC CAA TAT AGA CAC TTT
398 Ala Cys Cys Leu Glu Phe His Gly Ile Asn Gln Tyr Arg His Phe
399 65 70 75 80
401 GAT TAT GCT GGA TAT GCA ACA ACA GTT GAA GAA ACT AAA CAA GAC
402 Asp Tyr Ala Gly Tyr Ala Thr Thr Val Glu Glu Glu Thr Lys Gln Asp
403 85 90 95 100
405 ATC TTA GAC GTG CTA ATT CCG TGC TAC CTG TGT CAC AAA CCG CTG TGT
406 Ile Leu Asp Val Leu Ile Arg Cys Tyr Leu Cys His Lys Pro Leu Cys
407 100 105 110 115
409 GAA GTC GAA AAG GTC AAA CAT ATA CTA ACC AAG GCG CCG TTC ATA AAG
410 Glu Val Glu Lys Val Lys His Ile Leu Thr Lys Ala Arg Phe Ile Lys
411 115 120 125 130
413 CTA AAT TGT ACG TGG AAG GGT CGC TGC CTA CAC TGC TGG ACA ACA TGC
414 Leu Asn Cys Thr Trp Lys Gly Arg Cys Leu His Cys Trp Thr Thr Cys
415 130 135 140 145
417 ATG GAA GAC ATG TTA CCC AAG CTT CCA TGG GAA GTG GTG CCT GTA CAA
418 Met Glu Asp Met Leu Pro Lys Leu Pro Trp Glu Val Val Pro Val Gln

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/000,004

DATE: 08/02/2000
TIME: 13:46:21

Input Set : A:\Sabe-137.app
Output Set: N:\CRF3\08022000\I000004.raw

L:35 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:36 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]